Tanto a teoria proposta por Sapir-Whorf como o modelo de análise distribucional formulado por Bloomfield inserem-se na situação linguística específica dos Estados Unidos naquele início de século. Havia no continente americano cento e cinquenta famílias de *línguas ameríndias – o equivalente a aproximadamente mil línguas* – apresentadas sob a forma de material linguístico oral ainda não descrito, o que representava um grande problema para os administradores e etnólogos da época. A perspectiva antropológica presente nos postulados de Sapir-Whorf e a psicologia comportamental que influenciou as ideias de Bloomfield encontram terreno fértil nesse contexto particular.

Esse contexto, portanto, marcou o estruturalismo dos Estados Unidos, diferenciando-o da linguística europeia. Pode-se dizer que, enquanto Sapir foi o pioneiro, Bloomfield foi o consolidador da linguística naquele país, criando uma teoria mais bem delimitada do que os linguistas anteriores.

Exercícios

1) Comente a afirmativa saussuriana:

   "A língua é um sistema cujas partes podem e devem ser consideradas em sua solidariedade sincrônica" (Saussure, 1975).

2) Defina os conceitos de "língua" e "fala".
3) Um dos postulados de base da linguística estrutural é que o signo é arbitrário. Explique o que significa essa afirmação.
4) A linguística estrutural reconhece o princípio saussuriano de que *todo o mecanismo linguístico repousa sobre relações de dois tipos: sintagmáticas e paradigmáticas*. Explique tal princípio.
5) A afirmativa de que "a linguística tem por único e verdadeiro objeto a língua considerada em si mesma e por si mesma", que finaliza o texto do *Curso de linguística geral*, é fundamental para que possamos compreender os postulados de Saussure. Faça alguns comentários a respeito dessa questão.
6) Aponte três características da linguística descritiva norte-americana (distribucionalismo) que fazem dela uma vertente do estruturalismo saussuriano.

Notas

[1] A noção inicial era a de *sistema*, proposta por Saussure. A noção de estrutura se desenvolveu do termo saussuriano: tendo sido estabelecido que a língua constitui um sistema, cumpre estabelecer como se estrutura esse sistema.

[2] Essa característica se desenvolveu de modo mais forte na chamada Escola de Copenhague, sobretudo com Louis Hjelmslev. As chamadas escolas de Praga e de Genebra, desenvolvendo uma linha um pouco diferente, procuraram relacionar essa estrutura com a noção de "função".

[3] Saussure caracteriza as onomatopeias autênticas como aquelas que representam imitações aproximativas e já meio convencionais de certos ruídos, em oposição àquelas que impressionam por sua sonoridade sugestiva, como, por exemplo, tilintar, chover e piar.

[4] O estruturalismo europeu está representado principalmente pela linguística funcional desenvolvida pela Escola de Praga. As questões relacionadas a essa vertente são tratadas em capítulo específico.

Gerativismo

Eduardo Kenedy

Neste capítulo, apresentam-se em linhas gerais os principais aspectos que caracterizam a corrente de estudos linguísticos conhecida como *gerativismo*. Analisaremos a concepção de linguagem humana que norteia as pesquisas dessa corrente, bem como faremos uma exposição da maneira gerativista de observar, descrever e explicar os fatos das línguas naturais. Trata-se de uma visão geral, introdutória e simplificada, destinada ao estudante que conhece pouco ou nada sobre o gerativismo. Nas indicações bibliográficas, apresentadas no fim do livro, o leitor encontrará sugestões de leituras em português para prosseguir nos estudos sobre o assunto.

A faculdade da linguagem

A *linguística gerativa* – ou *gerativismo*, ou, ainda, *gramática gerativa* – é uma corrente de estudos da ciência da linguagem que teve início nos Estados Unidos, no final da década de 1950, a partir dos trabalhos do linguista Noam Chomsky, professor do Instituto de Tecnologia de Massachussets, o MIT. Considera-se o ano de 1957 a data do nascimento da linguística gerativa, ano em que Chomsky publicou seu primeiro livro, *Estruturas sintáticas*. Trata-se, portanto, de uma linha de pesquisa linguística que já possui cinquenta anos de plena atividade e produtividade. Ao longo desse meio século, o gerativismo passou por diversas modificações e reformulações, que refletem a preocupação dos pesquisadores dessa corrente em elaborar um modelo teórico formal, inspirado na matemática, capaz de descrever e explicar abstratamente o que é e como funciona a linguagem humana.

A linguística gerativa foi inicialmente formulada como uma espécie de resposta e rejeição ao modelo behaviorista de descrição dos fatos da linguagem, modelo esse que foi dominante na linguística e nas ciências de uma maneira geral durante toda

a primeira metade do século XX. Para os behavioristas, dentre os quais se destacava o linguista norte-americano Leonard Bloomfield, a linguagem humana era interpretada como um condicionamento social, uma resposta que o organismo humano produzia mediante os estímulos que recebia da interação social. Essa resposta, a partir da repetição constante e mecânica, seria convertida em hábitos, que caracterizariam o comportamento linguístico de um falante. Vejamos, por exemplo, como Bloomfield (1933: 29-30) descrevia a maneira pela qual uma criança aprendia a falar uma língua:

> Cada criança que nasce num grupo social adquire hábitos de fala e de resposta nos primeiros anos de sua vida. [...] Sob estimulação variada, a criança repete sons vocais. [...] Alguém, por exemplo, a mãe, produz, na presença da criança, um som que se assemelha a uma das sílabas de seu balbucio. Por exemplo, ela diz *doll* [boneca]. Quando esses sons chegam aos ouvidos da criança, seu hábito entra em jogo e ela produz a sílaba de balbucio mais próxima, *da*. Dizemos que nesse momento a criança começa a imitar. [...] A visão e o manuseio da boneca e a audição e a produção da palavra *doll* (isto é, *da*) ocorrem repetidas vezes em conjunto, até que a criança forma um hábito. [...] Ela tem agora o uso de uma palavra.

Para um behaviorista, a linguagem humana é exatamente o que descreveu Bloomfield: um fenômeno externo ao indivíduo, um sistema de hábitos gerado como resposta a estímulos e fixado pela repetição. Numa resenha feita em 1959 sobre o livro *Comportamento verbal*, escrito por B. F. Skinner, professor da famosa universidade de Harvard e principal teórico do behaviorismo, Chomsky apresentou uma radical e impiedosa crítica à visão comportamentalista da linguagem sustentada pelos behavioristas. Na resenha, Chomsky chamou a atenção para o fato de um indivíduo humano sempre agir criativamente no uso da linguagem, isto é, a todo momento, os seres humanos estão construindo frases novas e inéditas, ou seja, jamais ditas antes pelo próprio falante que as produziu ou por qualquer outro indivíduo.

Por isso, todos os falantes são criativos, desde os analfabetos até os autores dos clássicos da literatura, já que todos criam infinitamente frases novas, das mais simples e despretensiosas às mais elaboradas e eruditas. Pensemos, por exemplo, na frase que acabamos de produzir aqui mesmo neste texto. É muito provável que ela nunca tenha sido proferida exatamente da maneira como o fizemos, bem como jamais será dita novamente da mesma forma. Chomsky chegou a afirmar, inclusive, que a *criatividade* é o principal aspecto caracterizador do comportamento linguístico humano, aquilo que mais fundamentalmente distingue a linguagem humana dos sistemas de comunicação animal.

De acordo com esse pensamento de Chomsky, se considerarmos a criatividade a principal característica da linguagem humana, então devemos abandonar o modelo teórico e metodológico do behaviorismo, já que nele não há espaço para eventos criativos, pois, para linguistas como Bloomfield, o comportamento linguístico de um indivíduo deve ser interpretado como uma resposta completamente previsível a partir de um dado *estímulo, tal como é possível* prever que um cão começará a latir ao ouvir, por exemplo, o som de uma campainha caso tenha sido treinado para isso.[1] Se o behaviorismo deve ser abandonado, como de fato foi após a publicação da resenha de Chomsky, o gerativismo se apresenta como um modelo capaz de superá-lo e substituí-lo.

Com as suas ideias, Chomsky revitalizou a concepção *racionalista* dos estudos da linguagem, em oposição franca e direta à concepção *empiricista* de Skinner, Bloomfield e demais estruturalistas norte-americanos e europeus. Para Chomsky, a capacidade humana de falar e entender uma língua (pelo menos), isto é, o comportamento linguístico dos indivíduos, deve ser compreendida como o resultado de um dispositivo inato, uma capacidade genética e, portanto, interna ao organismo humano (e não completamente determinada pelo mundo exterior, como diziam os behavioristas), a qual deve estar radicada na biologia do cérebro/mente da espécie e é destinada a constituir a competência linguística de um falante. Essa disposição inata para a competência linguística é o que ficou conhecido como *faculdade da linguagem*.

Há, de fato, muitas evidências de que a linguagem seja uma faculdade natural à espécie humana. Pensemos, por exemplo, que, excluindo-se os casos patológicos graves, todos os indivíduos humanos, de todas as raças, em qualquer condição social, em todas as regiões do planeta e em todos os tempos da história foram e são capazes de manifestar, ao cabo de alguns anos de vida e sem receber instrução explícita para tanto, uma competência linguística – a capacidade natural e inconsciente de produzir e entender frases. É notável que nenhum outro ser do planeta, a não ser o próprio homem, seja capaz de dominar naturalmente um sistema de linguagem tão complexo como uma língua natural mesmo após muitos anos de treinamento. E nem mesmo o mais potente e arrojado dos computadores modernos é capaz de reproduzir artificialmente os aspectos mais elementares do comportamento linguístico de uma criança de menos de 3 anos de idade, como criar ou compreender uma frase completamente nova.

Não é por outra razão que a faculdade da linguagem é a característica mental mais marcante que separa os humanos dos demais primatas superiores e do resto do mundo natural. O papel do gerativismo no seio da linguística é constituir um modelo teórico capaz de descrever e explicar a natureza e o funcionamento dessa faculdade, o que significa procurar compreender um dos aspectos mais importantes da mente humana, como afirmou o próprio Chomsky (1980: 9):

> Uma das razões para estudar a linguagem (*exatamente a razão gerativista*) – e para mim, pessoalmente, a mais premente delas – é a possibilidade instigante de ver a linguagem como um "espelho do espírito", como diz a expressão tradicional. Com isto não quero apenas dizer que os conceitos expressados e as distinções desenvolvidas no uso normal da linguagem nos revelam os modelos do pensamento e o universo do "senso comum" construídos pela mente humana. Mais instigante ainda, pelo menos para mim, é a possibilidade de descobrir, através do estudo da linguagem, princípios abstratos que gover-nam sua estrutura e uso, princípios que são universais por necessidade biológica e não por simples acidente histórico, e que decorrem de características mentais da espécie humana.

Com o gerativismo, as línguas deixam de ser interpretadas como um comportamento socialmente condicionado e passam a ser analisadas como uma faculdade mental natural. A morada da linguagem passa a ser a mente humana.

O modelo teórico

Naturalmente, apenas postular a existência da faculdade da linguagem como um dispositivo inato que permite aos humanos desenvolver uma competência linguística não resolveria todos os problemas da linguística gerativa. Era (e ainda é) preciso descrever exatamente como é essa faculdade, como ela funciona e como é possível que ela seja geneticamente determinada se as línguas do mundo parecem tão diferentes entre si. Para dar conta dessa aparente contradição entre a hipótese da faculdade da linguagem e as milhares de línguas existentes no planeta, os linguistas da corrente gerativa vêm elaborando teorias que procuram explicar o funcionamento da linguagem na mente das pessoas. Ao observar os fatos das línguas naturais, um gerativista faz-se perguntas como:

- O que há em comum entre todas as línguas humanas e de que maneira elas diferem entre si?
- Em que consiste o conhecimento que um indivíduo possui quando é capaz de falar e compreender uma língua?
- Como o indivíduo adquire esse conhecimento?
- De que maneira esse conhecimento é posto em uso pelo indivíduo?
- Quais são as sustentações físicas presentes no cérebro/mente que esse conhecimento recebe?

Para procurar responder a perguntas como essas, a linguística gerativa propõe-se a analisar a linguagem humana de uma forma matemática e abstrata (formal), que se afasta bastante do trabalho empírico da gramática tradicional, da linguística estrutural e da sociolinguística, e se aproxima da linha interdisciplinar de estudos da mente humana conhecida como *ciências cognitivas*. A maneira pela qual tais perguntas vêm sendo respondidas constitui o *modelo teórico* do gerativismo.

Ao longo dos anos, linguistas de todas as partes do mundo (inclusive do Brasil, desde a década de 1970) têm trabalhado na formulação e no refinamento do modelo teórico gerativista. O mais importante deles é o próprio Chomsky, mas existem muitos estudiosos que dele discordam e acabam formalizando modelos alternativos, que às vezes divergem crucialmente do modelo chomskyano. Não há qualquer dúvida de que Chomsky seja não só o criador como, principalmente, o mais influente teórico da linguística gerativa – e um dos mais importantes estudiosos da linguagem de todos os tempos –, no entanto não se deve traçar um sinal de igual entre Chomsky e o gerativismo. É muito comum encontrarmos gerativistas que não são chomskyanos, apesar de que, quase sempre, ser chomskyano significa ser gerativista. Vejamos a seguir *as principais características* dos modelos chomskyanos (e convidamos o leitor, ao avançar em seus estudos, a conhecer os modelos diferentes).

A gramática como sistema de regras

A primeira elaboração do modelo gerativista ficou conhecida como *gramática transformacional* e foi desenvolvida e reformulada diversas vezes durante as décadas de 1960 e 1970. Os objetivos dessa fase do gerativismo consistiam em descrever como os constituintes das sentenças eram formados e como tais constituintes transformavam-se em outros por meio da aplicação de regras. Por exemplo, a sentença "o estudante leu o livro" possui cinco itens lexicais, que estão organizados entre si através de relações estruturais que chamamos de *marcadores sintagmáticos*, e tais marcadores poderiam sofrer regras de transformação de modo a formar outras sentenças, como "o livro foi lido pelo estudante", "o que o estudante leu?", "quem leu o livro?", etc. Ou seja, os gerativistas perceberam que as infinitas sentenças de uma língua eram formadas a partir da aplicação de um finito sistema de regras (a gramática) que transformava uma estrutura em outra (sentença ativa em sentença passiva, declarativa em interrogativa, afirmativa em negativa, etc.) – e é precisamente esse sistema de regras que, então, se assumia como o conhecimento linguístico existente na mente do falante de uma língua, o qual deveria ser descrito e explicado pelo linguista gerativista.

Vejamos um exemplo. A sentença (S) "o aluno leu o livro" é formada pela relação estrutural entre o sintagma nominal (SN) "o aluno" e o sintagma verbal (SV) "leu o livro". O SN é formado pelo determinante (DET) "o" e pelo nome (N) "aluno"; e o SV, por sua vez, é formado pelo verbo (V) "leu" e pelo outro SN "o livro", o qual se forma também por uma relação entre DET e N, no caso "o" e "livro" respectivamente. Toda essa estrutura sintagmática pode ser mais claramente visualizada no esquema abaixo, denominado *diagrama arbóreo* (ou, simplesmente, *árvore*), que é a famosa maneira pela qual os gerativistas representam estruturas sintáticas.
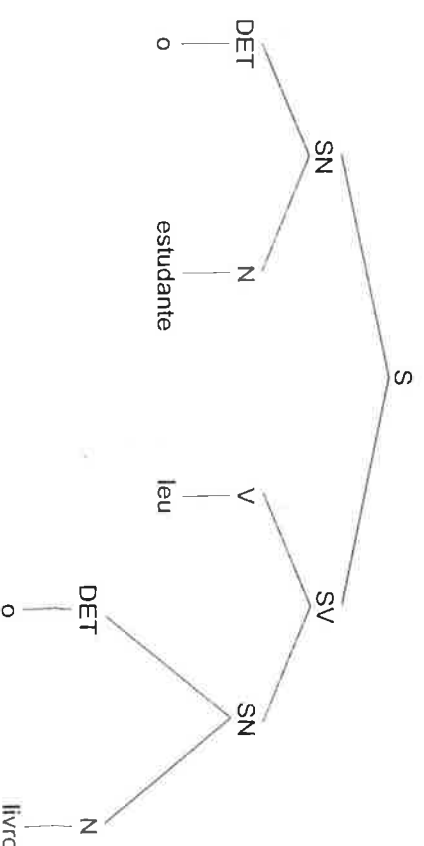
Figura 1: representação arbórea.

Essas regras de composição sintagmática explicam como uma estrutura simples como esta é gerada, mas não são suficientes para explicar como uma outra estrutura relacionada, como a voz passiva, seria formada a partir da estrutura de base, no caso, a voz ativa. Para dar conta da relação entre estruturas diferentes, mas relacionadas, os gerativistas formularam as regras transformacionais. Essencialmente, uma transformação forma uma estrutura a partir de uma outra previamente existente. A estrutura primeiramente formada é chamada de *estrutura profunda*, e a estrutura dela derivada chama-se *estrutura superficial*. Nesse sentido, a voz ativa é interpretada como a estrutura profunda sobre a qual são aplicadas as regras transformacionais que geram a voz passiva, a estrutura superficial.
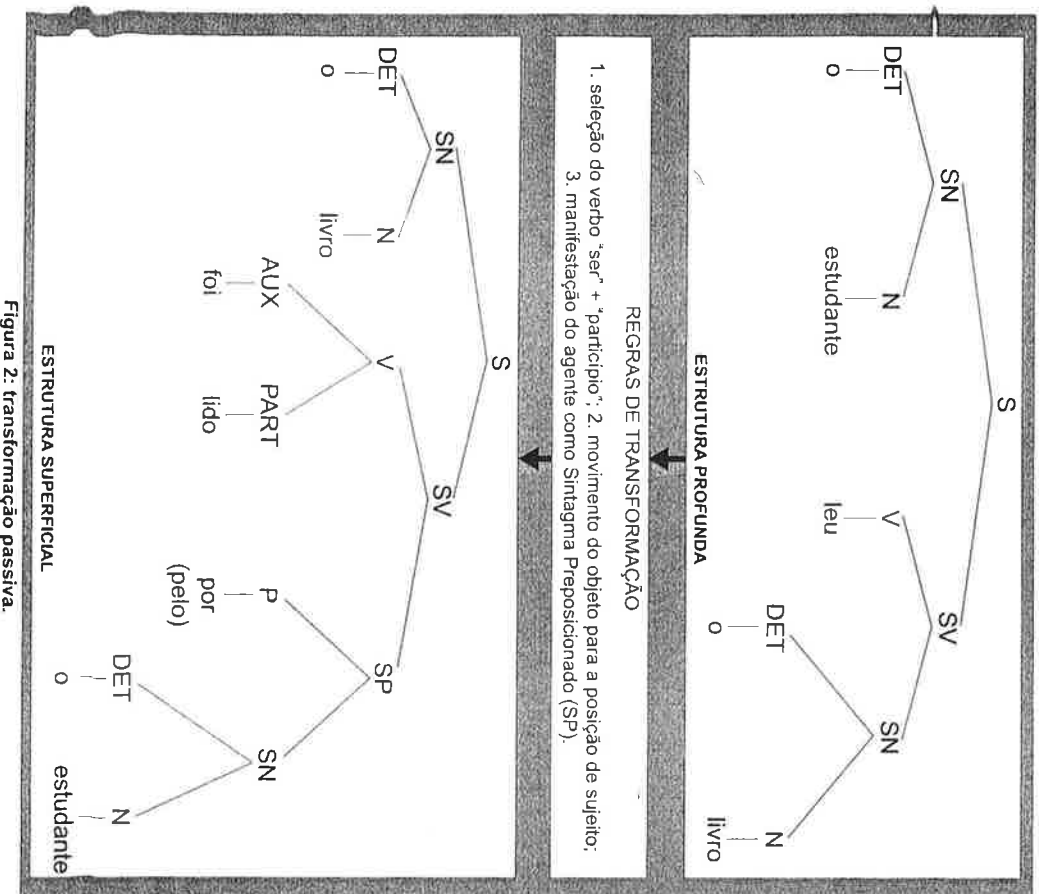
Figura 2: transformação passiva.

Na década de 1990, a ideia da transformação de uma estrutura profunda numa estrutura superficial seria abandonada em favor de uma visão que não mais *representava* estruturas, e sim as *derivava* – mostrando os passos pelos quais uma estrutura é formada (derivada) sem que ela tenha de ser comparada com uma outra estrutura independente. Não obstante, a ideia das transformações como operações computacionais (fenômenos sintáticos) que derivam sentenças é o tópico central da pesquisa gerativista até o presente momento.

Outro centro de atenção dos gerativistas sempre foi compreender como é possível que os falantes de uma língua tenham *intuições* sobre as estruturas sintáticas que produzem e ouvem. Por exemplo, todo falante nativo do português sabe que uma frase como "quantos livros você já escreveu?" é perfeitamente normal e pode ser falada por qualquer um de nós sem causar estranhamento. Trata-se, portanto, de uma frase *gramatical*, normal na língua. Esse mesmo falante do português também sabe, pela sua intuição, que uma frase como "*que livro você conhece uma pessoa que escreveu?" não é normal, é estranha, é uma frase *agramatical* da língua (e, por isso, aparece antecedida do asterisco, que indica a agramaticalidade).

Ora, como é que o falante sabe disso? Como ele consegue distinguir uma frase gramatical de uma frase agramatical em sua língua? Note bem: estamos falando de um conhecimento implícito, inconsciente e natural acerca da língua que todos os falantes nativos possuem, e não das regras da gramática normativa que aprendemos na escola. Na escola nunca são analisadas construções como a frase agramatical citada.

a) quantos livros você já escreveu? → gramatical
b) * que livro você conhece uma pessoa que escreveu → agramatical

Figura 3: gramaticalidade *vs.* agramaticalidade.

Um outro exemplo: na sentença "João disse que ele vai se casar", todo falante nativo de português sabe que o pronome "ele" pode referir-se tanto a João quanto a outra pessoa qualquer (do sexo masculino), diferente de João e citada anteriormente no discurso – isto é, a frase pode dizer que o próprio João vai se casar ou que um outro homem vai se casar. Mas na frase "Ele disse que João vai se casar" o falante sabe que "ele" não pode ser a mesma pessoa que "João" – e, nesse caso, a frase diz somente que João vai se casar. Todos os falantes de português conhecem inconscientemente essas pequenas regras que acabamos de descrever e é por isso que entendem e produzem as frases de sua língua. Mas como isso é possível? Como podemos saber essas coisas se ninguém nos ensina explicitamente como a língua funciona?

Esse conhecimento linguístico inconsciente que o falante possui sobre a sua língua e que lhe permite essas intuições é o que denominamos *competência linguística* – o conhecimento interno e tácito das regras que governam a formação das frases da língua. A competência linguística não é a mesma coisa que o comportamento linguístico do

indivíduo, aquelas frases que de fato uma pessoa pronuncia quando usa a língua. Esse uso concreto da língua denomina-se *desempenho linguístico* (também conhecido por *performance* ou, ainda, *atuação*) e envolve diversos tipos de habilidade que não são linguísticas, como atenção, memória, emoção, nível de estresse, conhecimento de mundo, etc. Imagine que você desejava pronunciar a frase "Vou tentar a sorte", mas enrolou a língua e acabou dizendo "vou tentar a torte". Ora, o que aconteceu foi apenas um erro de execução, com a preservação do segmento /t/ no início da palavra "sorte", o que não significa que seu conhecimento sobre o português tenha sido abalado. O que ocorreu não foi um problema de *conhecimento*, mas de *uso*, de *desempenho*, de *performance* da língua.

Classicamente o interesse central das pesquisas gerativistas recai na competência linguística dos falantes – muito embora só se possa ter acesso a ela através do desempenho –, pois é essa competência que torna o indivíduo capaz de falar e compreender uma língua. De acordo com essa abordagem, é somente através do estudo da competência que será possível elaborar uma teoria formal que explique o funcionamento abstrato da linguagem na mente dos indivíduos.

Em razão desse interesse central na competência linguística, os estudos clássicos do gerativismo não costumam usar dados linguísticos reais (*performance*) retirados do uso concreto da língua na vida cotidiana. O que interessa fundamentalmente ao gerativista é o funcionamento da mente que permite a geração das estruturas linguísticas observadas nos dados de qualquer *corpus* de fala, mas não lhe interessam esses dados em si mesmos ou em função de qualquer fator extralinguístico, como o contexto comunicativo ou as variáveis sociais que influenciam o uso da linguagem. Os gerativistas usam como dados para as suas análises principalmente (1) testes de gramaticalidade, nos quais frases são expostas a falantes nativos de uma língua, que devem utilizar sua intuição e distinguir as frases gramaticais das agramaticais, e (2) a intuição do próprio linguista, que, afinal, também é um falante nativo de sua própria língua.

Não obstante, os gerativistas que fazem pesquisas aplicadas (psicolinguistas, neurolinguistas, etc.)[2] também observam os dados do uso da língua, em situação natural ou em situação experimental, procurando extrair deles informações para o modelo de explicação da competência linguística. Por exemplo, esses gerativistas se interessam por (1) testes e experimentos psicolinguísticos, com pessoas de todas as idades, nos quais os informantes são levados a produzir ou interpretar determinados tipos de estruturas linguísticas; (2) testes e experimentos de aquisição da linguagem com crianças, além de gravações da fala natural destas; (3) testes e experimentos neurolinguísticos através dos quais se observa o funcionamento do cérebro quando em atividade linguística e também o desempenho linguístico de pacientes afásicos (pessoas que possuem dificuldades no desempenho linguístico em decorrência de uma lesão cerebral, na maior parte das vezes); (4) evidências das mudanças linguísticas por que passam as línguas, como uma maneira de compreender o que ocorre com

a gramática quando algum de seus componentes se transforma ao longo do tempo, perdendo ou ganhando formas. Esse último tipo de análise gerativista é o que mais se aproxima da linguística baseada em dados concretos do uso da língua (*corpus*). No Brasil, trabalharam e trabalham nessa linha, que ficou conhecida como sociolinguística paramétrica, linguistas de importância e reconhecimento internacional como Fernando Tarallo, Mary Kato, Maria Eugênia Duarte, entre outros.

A gramática universal: princípios e parâmetros

Com a evolução da linguística gerativa no início dos anos 1980, a ideia da competência linguística como um sistema de regras específicas cedeu lugar à hipótese da gramática universal (GU). Deve-se entender por GU o conjunto das propriedades gramaticais comuns compartilhadas por todas as línguas naturais, bem como as diferenças entre elas que são previsíveis segundo o leque de opções disponíveis na própria GU. A hipótese da GU representa um refinamento da noção de faculdade da linguagem sustentada pelo gerativismo desde o seu início: a faculdade da linguagem é o dispositivo inato, presente em todos os seres humanos como herança biológica, que nos fornece um algoritmo, isto é, um sistema gerativo, um conjunto de instruções passo a passo – como as inscritas num programa de computador – o qual nos torna aptos para desenvolver (ou adquirir) a gramática de uma língua. Esse algoritmo é a GU.

Para procurar descrever a natureza e o funcionamento da GU, os gerativistas formularam uma teoria chamada de *princípios e parâmetros*. Essa teoria possui pelo menos duas fases: a fase da teoria da regência e da ligação (TRL), que perdurou por toda a década de 1980, e o programa minimalista (PM), em desenvolvimento desde o início da década de 1990 até o presente.

As pesquisas da teoria de princípios e parâmetros foram e são desenvolvidas sobretudo na área da sintaxe, pois é exatamente nas estruturas sintáticas que mais evidentemente se percebem as grandes semelhanças entre todas as línguas do mundo, mesmo entre aquelas que não possuem nenhum parentesco, o que facilita o estudo da GU. Por exemplo, todas as línguas do mundo possuem estruturas como orações adjetivas, orações interrogativas e funções sintáticas como sujeito, predicado, complementos.

A possibilidade de estudar a sintaxe isolada dos demais componentes da gramática (léxico, fonologia, morfologia, semântica) é consequência de um conceito fundamental do gerativismo, o de *gramática modular*. Segundo ele, os componentes da gramática devem ser analisados como módulos autônomos, independentes entre si, no sentido de que são governados por suas próprias regras e não sofrem influência direta dos outros módulos. Isto é, o funcionamento de um módulo como, digamos, a sintaxe, é cego em relação às operações da fonologia, por exemplo. Naturalmente existem pontos de interseção entre os módulos da gramática, afinal a sintaxe cria sintagmas e sentenças a partir das palavras

do léxico, e o produto final da sintaxe (a sentença) deve receber uma leitura fonológica e também uma interpretação semântica básica, que no gerativismo se chama forma lógica. Podemos visualizar essa interação entre os módulos da gramática no esquema a seguir.
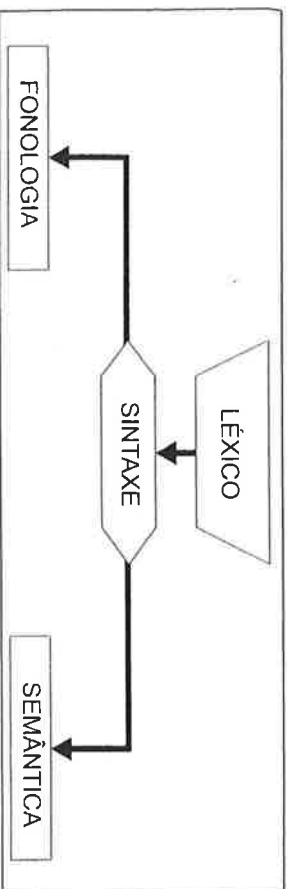
Figura 4: o modelo de gramática.

Nessa ilustração, vemos que o elemento central da gramática é a sintaxe. Ela retira do léxico as palavras com as quais construirá, segundo suas próprias regras, estruturas como sintagmas e sentenças, que da sintaxe são encaminhadas à preparação para a pronúncia, no módulo fonológico, e para a interpretação formal, no módulo semântico. Nessa maneira de compreender o funcionamento da gramática, a morfologia é interpretada como parte do léxico, já que dá conta da estrutura interna da palavra, e também como parte da fonologia, uma vez que deve dar conta das alterações mórficas fonologicamente condicionadas.

No programa minimalista atual, entendemos por "princípio" as propriedades gramaticais que são válidas para todas as línguas naturais, ao passo que "parâmetro" deve ser compreendido como as possibilidades (limitadas sempre de maneira binária) de variação entre as línguas. Por exemplo, quando analisamos as sentenças (a) "João disse que ele vai se casar" e (b) "Ele disse que João vai se casar", vimos que em (a) o pronome "ele" pode referir-se tanto a "João" quanto a qualquer outro homem anteriormente citado no discurso, mas na frase (b) "ele" não pode se referir a "João" e necessariamente faz referência a um outro homem.

Essa diferenciação entre a referencialidade do pronome "ele" nas duas frases pode ser explicada da seguinte maneira: nesse contexto, o pronome faz referência a algum elemento que precisa ter sido citado anteriormente no texto – trata-se de um pronome anafórico. É um princípio da GU que uma anáfora necessariamente deve suceder o seu referente, e nunca o contrário. É por isso que na frase (a) "ele" pode ser tanto "João" quanto outro homem citado numa frase anterior, já que ambos os termos antecedem o pronome. Já no caso de (b) "João" não pode ser o referente de "ele", *pois o pronome* antecede o nome. Se traduzíssemos (a) e (b) para qualquer língua do mundo, o resultado seria sempre o mesmo: em (b) seria impossível ligar o pronome ao nome citado, mas em (a) isso pode ocorrer. Trata-se, portanto, de um princípio da GU, exatamente igual em todas as línguas naturais.

Vejamos agora um exemplo de parâmetro. Se considerarmos que o valor semântico básico da frase (a) seja, digamos, algo como "João disse que ele mesmo, o próprio João, vai se casar", saberemos que "ele" se refere a "João". "João" é o sujeito da oração principal, e "ele" é o sujeito da oração subordinada. Dizemos, então, que os sujeitos das duas orações são correferenciais. O que é interessante nesse exemplo é que o sujeito da segunda oração poderia não ser preenchido por um pronome anafórico, isto é, o sujeito da oração subordinada poderia ser oculto – que na linguística gerativa chamamos tecnicamente de sujeito nulo (representado aqui informalmente por Ø) –, como ocorre na sentença (c) "João disse que Ø vai se casar".

i) João disse que ele vai se casar ("ele" → sujeito preenchido)
ii') João disse que Ø vai se casar ("Ø" → sujeito nulo)

Podemos dizer que a língua portuguesa se caracteriza por suportar a ocorrência de sujeitos nulos, como ocorre também nessas frases "Ø saí ontem", "Ø fomos ao cinema", "Ø fez o trabalho?", "Ø choveu ontem", etc. Tanto nesses casos quanto na oração subordinada em (ii), o SN sujeito do SV predicado não possui nenhum elemento pronunciado, está vazio, nulo, como se ilustra na sentença (S) abaixo.
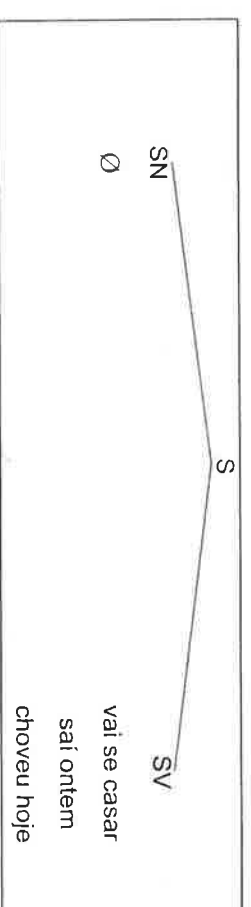
Figura 5: parâmetro do sujeito nulo (+).

Deixar o sujeito nulo é uma propriedade do português e também de outras línguas, como o espanhol, o italiano, mas essa propriedade não é comum a todas as línguas humanas. Se traduzíssemos as sentenças do quadro acima para línguas como o inglês e o francês, teríamos necessariamente de preencher o SN sujeito com um elemento pronominal, pois nessas línguas o sujeito nulo é uma estrutura agramatical. A frase (ii), por exemplo, só poderia apresentar, em inglês, o pronome anafórico *he*, e nunca o sujeito nulo "Ø", independentemente da referencialidade da anáfora pronominal ou zero:

(i) John said that he is going to get married
(ii) * John said that Ø is going to get married

A existência de sujeitos nas sentenças é um princípio da GU, mas a possibilidade de deixá-los nulos nas frases é um parâmetro da GU, pois línguas como o português se caracterizam como [+ sujeito nulo], enquanto línguas como o inglês são [– sujeito nulo]. É por essa razão que dissemos que os parâmetros que diferenciam as línguas são previsíveis e distribuídos sempre de maneira binária (+ ou – o parâmetro x). O léxico, por exemplo,

não é um fator de diferenciação entre as línguas que possa ser interpretado como opção paramétrica, já que o léxico é sempre arbitrário e, por isso mesmo, imprevisível.
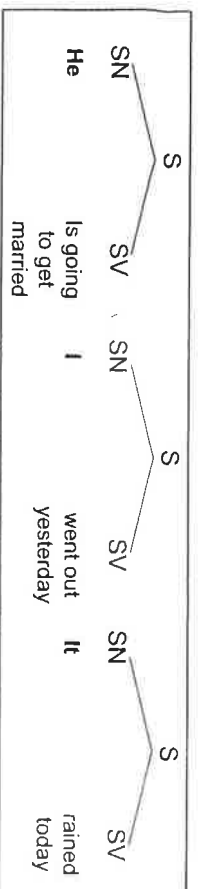
Figura 6: parâmetro do sujeito nulo (–).

Ao compararmos as figuras 5 e 6, percebemos que somente línguas como o português (e também o espanhol, o italiano, etc.) permitem o sujeito nulo "Ø", casos que conhecemos pela gramática tradicional como *sujeito oculto, indeterminado e inexistente*. Como indica a figura 6, línguas como o inglês (bem como o francês, o alemão, etc.) não permitem o sujeito nulo e exigem o preenchimento do SN sujeito da frase nem que seja com um pronome expletivo (sem conteúdo semântico), como o *it* do inglês.

O projeto da linguística gerativa é observar comparativamente as línguas humanas – com os seus milhares de fenômenos morfofonológicos, sintáticos, semânticos e sua suntuosa complexidade – com o objetivo de descrever os princípios e os parâmetros da GU que subjazem à competência linguística dos falantes, para, assim, poder explicar como é a faculdade da linguagem, essa parte notável da capacidade mental humana.

O FOXP2 e a genética da linguagem

Em outubro de 2001, um geneticista inglês chamado Anthony Monaco, professor da Universidade de Oxford e integrante do Projeto Genoma Humano, anunciou a descoberta do primeiro gene que aparentemente está destinado a controlar a capacidade linguística humana: o FOXP2. Monaco estudou diversas gerações da família K. E.,[3] e constatou que todos os seus membros possuíam distúrbios de linguagem que não estavam associados a algum problema físico superficial como língua presa, audição ineficiente, etc.

Esses distúrbios diziam respeito à conjugação verbal, à distribuição e à referencialidade dos pronomes, à elaboração de estruturas sintáticas complexas, como as orações subordinadas. O interessante é que os avós, pais, filhos e netos da família K. E. não possuíam aparentemente nenhum outro distúrbio cognitivo além desses problemas com o sistema linguístico. Monaco analisou amostras de DNA dessa família e descobriu que uma única unidade de DNA de um só gene estava corrompida. O FOXP2 é um dos setenta genes diferentes que compõem o cromossomo 7, que é responsável pela arquitetura genética do cérebro humano. Esse gene, o FOXP2, possui 2.500 unidades de DNA, e só uma delas apresenta problemas na genética da família K. E. Monaco estava quase certo de que esse gene deveria ser responsável pela capacidade genética associada à linguagem, e teve certeza disso quando descobriu o jovem inglês C. S., que não possuía parentesco com os K. E., mas apresentava os mesmos distúrbios linguísticos que os membros dessa família. Monaco analisou o FOXP2 de C. S. e constatou o que presumia: C. S. apresentava um defeito na mesma unidade de DNA do FOXP2 deficiente da família K. E. Daí o geneticista proclamou o que pode ser a descoberta do primeiro gene responsável pela genética da linguagem humana.

Independentemente de as pesquisas de Anthony Monaco serem confirmadas ou não – e há muitos geneticistas que as refutam –, o importante é que elas abriram ou aprofundaram a discussão, fora do âmbito da linguística gerativa, sobre as bases genéticas da linguagem humana. O FOXP2 é um gene existente também em outros primatas, como chimpanzé e gorilas, mas em quantidade muito reduzida – e isso pode explicar a limitada capacidade de comunicação linguística desses animais.

De fato, se o mapeamento dos genes humanos apontar, como a hipótese FOXP2 esboça, a existência de genes cuja função na genética de nossa espécie é controlar o uso de pronomes, a construção de orações subordinadas, a flexão de verbos, etc., então a faculdade da linguagem e sua disposição na GU através de princípios e parâmetros podem passar a ser considerados não mais hipóteses abstratas mas sim fatos do mundo natural. Consequentemente, a linguística gerativa será a corrente da ciência da linguagem que travará forte diálogo com as ciências naturais.

Exercícios

1) Em seu livro *O instinto da linguagem*, o linguista e psicólogo norte-americano Steven Pinker afirmou que a linguagem natural é um instinto da espécie humana, uma capacidade que herdamos da natureza. Para Pinker, assim como as aranhas são naturalmente programadas para tecer teias, os humanos são programados para falar (pelo menos) uma língua. Explique por que essa afirmação de Pinker deve ser considerada coerente com os fundamentos da linguística gerativa. Você concorda, em parte ou completamente, com a afirmação do psicólogo-linguista? Vê nela algum exagero? Comente.
2) Leia as sentenças abaixo. Ponha um asterisco antes daquelas que considerar, segundo a sua intuição, agramaticais e escreva OK depois daquelas que considerar gramaticais. Logo após explique: de onde vem essa intuição sobre as frases da língua?
    a) Parece que os alunos estão cansados.
    b) Os alunos, parece que estão cansados.
    c) Os alunos parecem estar cansados.
    d) Os alunos parecem estarem cansados.
    e) Parece os alunos estarem cansados.
    f) Parece os alunos estar cansados.
    g) Os alunos, parece que eles estão cansados.
3) No final da festa do aniversário do seu filho, dona Maria ia anunciar que estava na hora de cortar o bolo e disse a seguinte frase: "Vamos, gente, está na hora de *b*ortar o *c*olo!". A própria falante riu do que disse e corrigiu a frase logo depois. O erro linguístico que dona Maria cometeu deve ser explicado como um problema na *competência* ou no *desempenho* linguístico? Explique.